

# Dynamical downscaling of ERA-40 in complex terrain using the WRF regional climate model

U. Heikkilä · A. Sandvik · A. Sorteberg

Received: 16 June 2010 / Accepted: 8 October 2010

© The Author(s) 2010. This article is published with open access at Springerlink.com

**Abstract** Results from a first-time employment of the WRF regional climate model to climatological simulations in Europe are presented. The ERA-40 reanalysis (resolution 1°) has been downscaled to a horizontal resolution of 30 and 10 km for the period of 1961–1990. This model setup includes the whole North Atlantic in the 30 km domain and spectral nudging is used to keep the large scales consistent with the driving ERA-40 reanalysis. The model results are compared against an extensive observational network of surface variables in complex terrain in Norway. The comparison shows that the WRF model is able to add significant detail to the representation of precipitation and 2-m temperature of the ERA-40 reanalysis. Especially the geographical distribution, wet day frequency and extreme values of precipitation are highly improved due to the better representation of the orography. Refining the resolution from 30 to 10 km further increases the skill of the model, especially in case of precipitation. Our results indicate that the use of 10-km resolution is advantageous for producing regional future climate projections. Use of a large domain and spectral nudging seems to be useful in reproducing the extreme precipitation events due to the better resolved synoptic scale features over the North

Atlantic, and also helps to reduce the large regional temperature biases over Norway. This study presents a high-resolution, high-quality climatological data set useful for reference climate impact studies.

**Keywords** WRF · Downscaling · ERA-40 · ENSEMBLES · Regional climate modelling

## 1 Introduction

A number of model studies have addressed the regional effects of future climate change in Europe (ENSEMBLES members 2009). These studies point to increased precipitation in Northern Europe in the future which can have important impacts on hydrology, vegetation and infrastructure in the influenced areas. In order to deliver reliable information for the society on these issues we need to focus on a local level. This study is a first step towards a higher resolution assessment of future climate prediction in Norway. We provide high-resolution climate parameters for Norway which are increasingly required and of crucial importance for driving various climate impact models. This study employs a new model for Europe (WRF: Weather Research and Forecasting, <http://www.wrf-model>). It provides a first-time comprehensive model evaluation for new users wanting to apply the WRF model for climatological simulations over Europe, and also a qualitative comparison of the performance of the WRF model against other state-of-the-art regional climate models.

A common approach used in regional climate simulations for this region has been to include only the continent of Europe with little ocean into the high-resolved regional model domain. In this study, we aim to improve the representation of climate in Europe by increasing the size of

---

U. Heikkilä (✉)  
Uni Bjerknes Centre, Bjerknes Centre for Climate Research,  
Bergen, Norway  
e-mail: [ulla.heikkila@uni.no](mailto:ulla.heikkila@uni.no)

A. Sandvik  
Institute for Marine Research (IMR),  
Bjerknes Centre for Climate Research,  
Bergen, Norway

A. Sorteberg  
Geophysical Institute, University of Bergen, Bergen, Norway

the regional model domain to cover the whole North Atlantic. Such a setup will increase the independency of the regional climate model from the driving data. In this way also the synoptic scale features on open water will be better resolved on our high-resolution domain (30 km) than in the driving ERA-40 reanalysis (1.1°) before they reach the coast of Europe. We apply the spectral nudging procedure to force the regional model to keep its large scale circulation consistent with the driving reanalysis data.

Norway is a country with a comparably complex topography (see Fig. 1). Its continental part stretches between 60°N and 70°N and is some 500 km wide in the south and less than 100 km at the narrowest in the central parts. It has a long coastline including narrow fjords and mountains which cover almost the whole country. The highest mountains reach an altitude of ~2,500 m. The western coast of Norway is subject to mild coastal climate with high amounts of precipitation, typically exceeding 2,000 mm/year. The precipitation on the coast is mainly large scale frontal precipitation driven by the low pressure systems in the North Atlantic. In the mountains there is a significant orographic enhancement of precipitation (e.g. Leung and Ghan 1995). The northern and the eastern parts, on the lee side of the mountains, are more continental and experience as little as 500–1,000 mm of precipitation a year. In contrast to the western part the extremes are usually connected to convective systems.

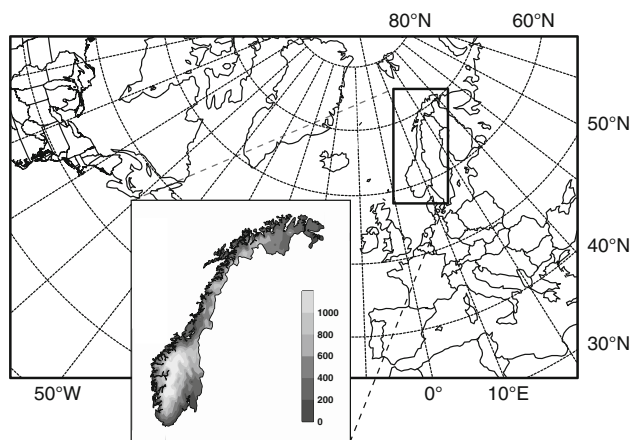
The focus of this study is on the validation of the WRF model in a period (1961–1990) for which many observations and model runs exist and on finding an optimal setup for future prediction simulations. The main questions to address are (1) how well does the WRF model agree with observations when run with “ideal” boundary conditions (ERA-40 reanalysis) and spectral nudging on the 30 and

10 km horizontal resolution? (2) Does the model simulation improve the driving ERA-40 reanalysis? Is the 10 km resolution adding significant value to the 30 km resolution? (3) How well does the WRF model capture the regional differences of climate in Norway? and (4) is the model able to reproduce the observed extreme values of precipitation and temperature? In order to put our model simulation into a larger context of the state-of-the-art of regional climate modelling we perform a comparison with 12 European and Canadian models which participated in the recently finished ENSEMBLES project (see Sect. 3.2 for more details on the project). The WRF and the ENSEMBLES model simulations are not directly comparable because of the different setup used. Keeping this in mind, the comparison presented in this study should be understood as qualitative.

## 2 Model description and setup

The model employed for this study is the WRF regional climate model (version 3.1.1). The WRF model has a rapidly growing user community and has been used for climatological studies, various case studies and operational weather forecasting among other purposes in the recent years. For the experiments of this study we used a large model domain of a size of 9,090 km (W–E)  $\times$  5,490 km (S–N) with a horizontal grid resolution of 30 km (Fig. 1). We included one nest inside the model domain with a horizontal resolution of 10 km. This nest has a size of 880 km (W–E)  $\times$  1,840 km (S–N). Both domains have 40 vertical levels reaching up to 50 hPa. The first reason for choosing such a large domain was that the precipitation in the western coast of Norway is mainly large scale and the moisture can have its origin far in the southwestern North Atlantic (Stohl et al. 2008). Another reason was that we wanted to find an optimal setup for subsequent future climate predictions. A larger model domain will give the regional model more freedom to develop its own synoptic and mesoscale circulation. This may be an advantage in regions where the climate change signal is strongly influenced by advective processes.

It has been noted that using a large domain may lead to deviation of the large scale features from the driving fields creating problems close to the boundaries (Jones et al. 1995; Koltzow et al. 2008). To reduce this risk, we use spectral nudging. Spectral nudging is a method which allows the passing of the driving global model information not only onto the lateral boundaries but also into the interior of the regional model domain (Waldron et al. 1996). The value of spectral nudging has been discussed in the literature (e.g. Alexandru et al. 2008; Míguez-Macho et al. 2004, 2005; Radu et al. 2008; Von Storch et al. 2000; Zahn et al. 2008) and there is some controversy. Most studies



**Fig. 1** The model domain used in this study: the larger domain with 30-km resolution covering the whole North Atlantic and the 10-km resolution nest covering Norway. The inset shows the terrain height in the 10-km nest (m)

agree that nudging too strongly will not allow the regional model to deviate much from the driving fields. While spectral nudging seems to reduce the sensitivity to the chosen model domain or grid size (Alexandru et al. 2008; Miguez-Macho et al. 2004) other studies show that it can affect extreme precipitation or high frequency dynamical phenomena (Alexandru et al. 2008; Radu et al. 2008). We conducted several tests to evaluate the sensitivity of the modelled surface variables to nudging. We found that spectral nudging has an important effect in keeping the large scale circulation of the regional model in phase with the global model, but does not constrain the model's ability to develop small scale features. The extreme precipitation events were actually better reproduced by the nudged run than the free run.

We applied the spectral nudging technique following previous studies by Miguez-Macho et al. 2005 and Radu et al. 2008. We nudged only in the outer domain in order to let the regional model create its own structures in the high-resolution nest. For the same reason we applied nudging only on vertical levels above the boundary layer. The threshold for wavelengths over which the waves were nudged was 1,000 km. Following Miguez-Macho et al. 2005 and Radu et al. 2008, the nudging was applied to *u* and *v* winds, temperature and geopotential height but not to humidity. The sensitivity to the strength of the nudging was tested but no significant differences were found between stronger (every 6 h) and weaker (every 24 h) nudging. We chose the weaker nudging approach in order to maximize the freedom of the regional model to deviate from the driving global fields.

We simulated the years from 1960 to 1990 because many climatology simulations exist for this period, such as the regional model runs of the EU-project ENSEMBLES (see Sect. 3.2). The first year was used to spin up the soil moisture and not included in the analysis. The driving global data used was the ERA-40 reanalysis (Uppala et al. 2005) with 1.1-degree horizontal resolution and 24 vertical pressure levels. The experiment was performed using the default setup of the WRF model for the physical parameterizations as much as possible to keep the runtime low. The cloud microphysical scheme used was the 3-class scheme (Hong et al. 2004), the Kain–Fritsch scheme (Kain 2004) for the convective parameterization, the Yonsei University (YSU) (Hong et al. 2006) planetary boundary layer scheme, the Monin–Obukhov scheme for surface layer processes and the 4-layer Noah land-surface model (Ek et al. 2003) for the land-surface and soil processes. We used the new MODIS land use data set to describe the vegetation and land use classes in Norway (<http://modis.gsfc.nasa.gov/>). The Community Atmosphere Model (CAM) schemes were used for short-wave and long-wave radiation (Collins et al. 2006). We tested the sensitivity of

the model to different microphysical schemes and found no significant differences between the simpler and the more sophisticated schemes on the spacial scales (10 km) or time scales (daily) of this study. We used the so called 1-way nesting procedure which passes information only from the outer domain to the inner nest. This is a common approach in climatological studies because of possible stability problems introduced by 2-way nesting.

### 3 Results

#### 3.1 Methods

The results obtained within this study are evaluated against daily surface observations (precipitation, 2-m temperature and 10-m wind speed) from the Norwegian meteorological office in a similar manner with Barstad et al. 2009. The observational network consists of several hundred meteorological stations covering the whole country and provides the best data available for Norway. The data was checked for continuity and consistency and only stations which contained a continuous 30-year data set were taken into account in this comparison. This left us with 316 stations of precipitation, 66 stations of 2-m temperature and 67 stations of 10-m wind speed data. The comparison was made using the nearest gridpoint of the model to the observations. Although the horizontal resolution of the model is quite high (30 and 10 km) the error of the elevation of the model gridpoint to the actual elevation can be large at some points, especially on the coast and in the mountain slopes. For temperature we used a simple lapse-rate correction assuming that the temperature drops 6 K each 1,000 m, as has been used in several studies (e.g. Barstad et al. 2009; Kostopoulou 2009). Assuming a constant negative lapse rate neglects many effects, such as the complexity of the temperature profile in a boundary layer. In a case of a winter-time inversion, for example, this correction actually increases the error. Still, without this correction the temperature bias will reflect mostly the smoothed topography and not the correctness of the model dynamics or the physical parameterizations. Moreover, comparison of the temperatures of the ERA-40 and both WRF simulations with very different resolutions would not be fair without such a correction.

In the case of wind the issue is more complicated as there is no standard procedure to correct for the altitude error. We know that the stations measuring wind in the mountains are located in small valleys which are not resolved by the model topography. Therefore, the wind observations are not necessarily representative for the areas they are located in. The wind observations are made optically which can introduce an error in some cases. In order to use only quality-checked

data representative for the location in question we use the ten coastal stations chosen by Barstad et al. 2009. These stations were chosen as recommended by the meteorological office responsible for the observations. Data was written out from the model every 6 h for the 30-km domain and every 3 h for the 10-km domain and daily means were calculated from these values.

### 3.2 Comparison with models participating in the ENSEMBLES project

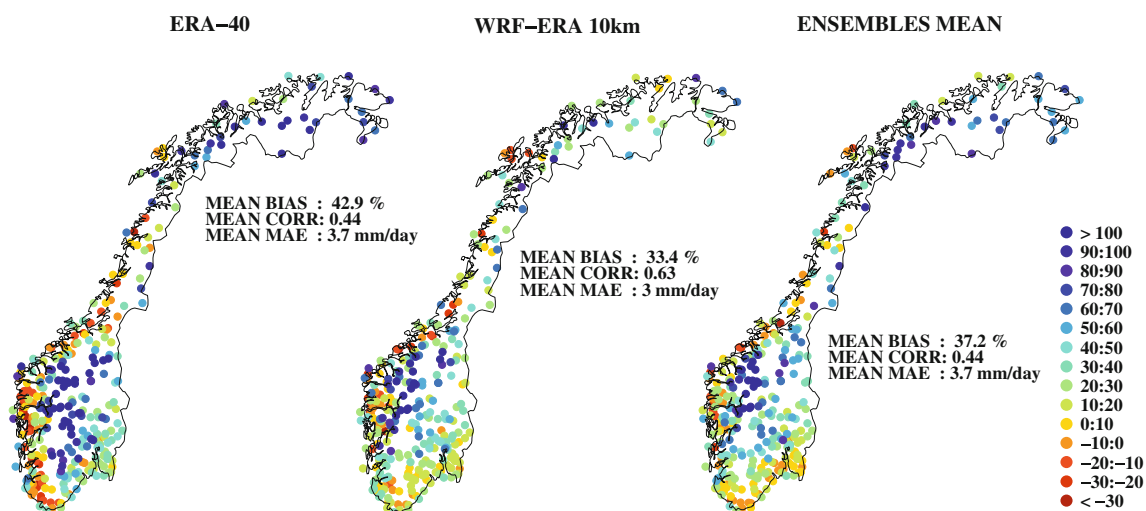
There has been large regional climate modelling and model inter-comparison activity in Europe during the recent years. The ENSEMBLES project (Ensembles-based predictions of climate changes and their impacts) was finished in the end of 2009 (ENSEMBLES members 2009). Its aim was to produce an ensemble of downscaled global future climate projections in order to provide the European society and economy with more detailed information on the future climate. Some ten state-of-the-art European and Canadian climate models took part in the project and several experiments were performed with different combinations of global model, greenhouse gas emission scenarios and horizontal resolution of the regional models. One part of it was, similar to the goal of this study, to validate the models driven with the ERA-40 reanalysis data for the period of 1958–2002. The results of this project give us an excellent opportunity to put our model results in a larger perspective and investigate how well the WRF model is performing within the spread of the ENSEMBLES models. We chose a set of 12 simulations with different models for comparison and performed the same analysis as with our simulations for the period of 1961–1990. These models are listed in Fig. 12.

We chose the 25-km resolution of the ENSEMBLES model runs to allow for a comparison as accurate as possible with our 30 and 10-km simulations. The number of vertical levels in the ENSEMBLES runs was lower than in our runs and varied from 19 to 32. No spectral nudging was used in these runs. Their domain size was smaller, covering Europe including the Mediterranean in the south but just only including the northernmost part of Norway in the north and not the whole Atlantic ocean. The analyzed precipitation, 2-m temperature and 10-m wind speed are daily means. The ENSEMBLES means shown are calculated as simple averages in each case and are not weighted based on model performance or any other way.

### 3.3 Precipitation

#### 3.3.1 Geographical distribution of precipitation bias

Figure 2 illustrates the bias of the total accumulated precipitation in the 30-year period for each precipitation station of the ERA-40 reanalysis, the 10-km WRF simulation and the mean of the ENSEMBLES models. The bias is calculated as a difference in percent between the modelled and observed 30-year total accumulated precipitation for each station separately and then averaged. The WRF 30 km simulation (not shown) performs very similarly to the 10 km one producing a slightly reduced mean bias (29.7%). Figure 2 also shows the mean statistics: the mean bias, the mean correlation coefficient between the observations and the model and the mean absolute error (MAE) calculated for the daily mean values of each station separately. We see that the 10-km WRF simulation performs similarly with the ENSEMBLES mean. The mean bias is



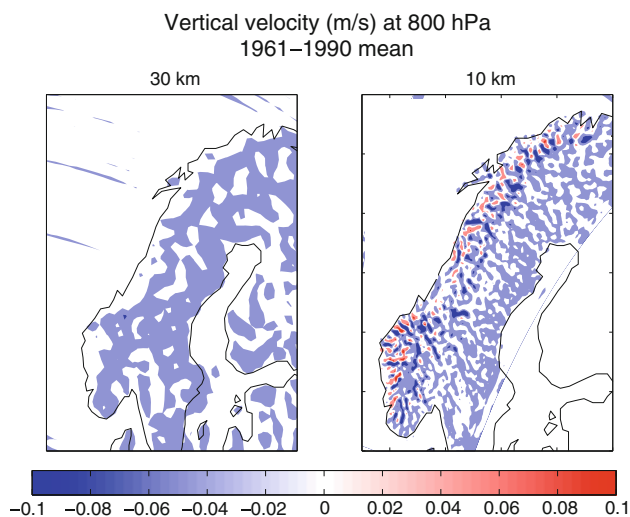
**Fig. 2** The 30-year total precipitation bias of the ERA-40 reanalysis, the WRF model (10 km) and the 12 model mean of the ENSEMBLES project. The bias is defined as the average deviation of the simulated

30 year accumulated daily mean precipitation of observations. The mean bias, mean correlation coefficient of the daily mean precipitation values, and the mean absolute error are also shown



comparable to the ENSEMBLES mean (33.4 and 37.2%, respectively). The main difference is the better correlation coefficient of the WRF simulation (0.63) than the ENSEMBLES mean (0.44) which is similar to the driving ERA-40 data (0.44). The correlation coefficient reflects the phase of the precipitation events. The phase of the precipitation events in the WRF simulations is good due to the spectral nudging procedure used which keeps the low pressure systems in phase with the ERA-40. The improved correlation of the WRF simulation from the reanalysis is probably caused by the higher horizontal resolution (30 km) on the North Atlantic which improves the representation of synoptic scale lows and even includes meso-scale features which are missing in the coarse ERA-40 data. Also the better resolved coastline and topography may improve the correlation.

Looking at the distribution of the bias we see that the coastal precipitation is generally well simulated. The bias is largest towards the inland on the lee side of the mountains. The elevation in the models is too low even at the 10–25 km resolution so that the orographic increase of the precipitation is too small and too much of the precipitation falls on the lee side of the mountains. The distribution of the 1961–1990 mean vertical velocity (Fig. 3) in the 30 and 10 km WRF simulations indicates that the orographic lifting is better resolved in the 10 km nest but even further refinement of the resolution would be needed to correct for this error. The bias is also large in the northern part of Norway which is a very dry region, but reduced by 20% compared to the ERA-40 or ENSEMBLES mean. The bias does not vary much seasonally (not shown), being slightly larger in percent during the driest period of the year (MAM) and lowest during the wettest months (SON).



**Fig. 3** The 1961–1990 mean of the vertical velocity (m/s) near the surface (800 hPa) of the 30-km and the 10-km WRF simulations

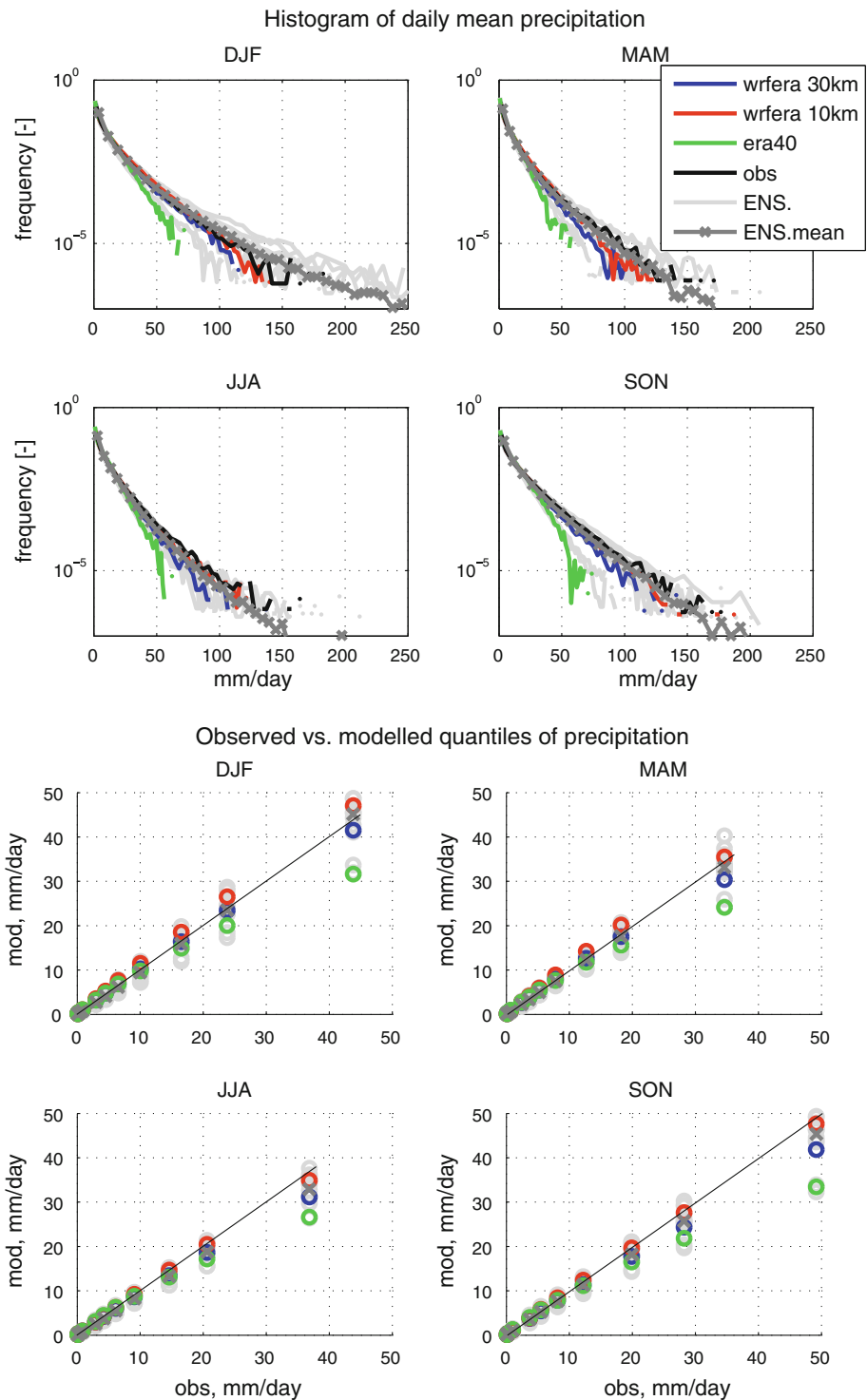
There is no large variability between the different regions. The bias is reduced in the WRF simulations, compared with ERA-40, during all seasons. The correlation coefficients are highest during the winter months (0.69) and lowest during the summer months (0.52). The reduced correlation during summer is caused by the more small scale convective precipitation whose phase does not profit from the spectral nudging procedure.

### 3.3.2 Histogram of the daily mean precipitation

One important measure for the skill of a model is its capability to simulate the intensity and frequency of individual precipitation events correctly. This can be assessed by looking at the distributions of individual events. Figure 4 shows the histogram of the daily mean precipitation modelled with WRF and compared with the ERA-40 reanalysis and the observations for the four seasons. The grey lines show the individual ENSEMBLES models and the dark grey line the ENSEMBLES mean. The precipitation of the ERA-40 reanalysis is correctly producing the lower end of the spectrum but the largest values of daily mean precipitation are completely missing. This is probably due to the low resolution of the reanalysis which does not resolve the orography in Norway and smoothens out the extreme events. The 30-km WRF simulation improves the representation of the extreme precipitation events ( $>50$  mm/day) significantly but cannot reproduce the highest extremes. Here, we see clear value added by further refining the resolution to 10 km, as many more of the observed extreme events are produced by the 10-km simulation due to the better resolved orographic lifting discussed in the previous section. There are no significant differences in the model performance between the seasons. Generally the agreement between the observed and modelled histograms is very good and improved compared with the ERA-40 data. The spread of the individual ENSEMBLES models is large. Some of the models hardly increase the number of extreme events from those of the ERA-40 reanalysis whereas other models highly overestimate the whole range of precipitation extremes, mainly during the winter (DJF) months. Some of the models produced unrealistic values exceeding 350 mm/day but the x-axis of the graph is truncated. The ENSEMBLES mean was calculated by pooling all daily mean values together. It reproduces the shape of the histogram very well but has too long a tail caused by the large overestimations of the extremes by a few models.

In order to ignore the few exaggerated extremes we look at the quantiles of the daily mean precipitation (lower panel of Fig. 4). The ERA-40 reanalysis again lacks the highest values of the spectrum and a few of the ENSEMBLES models perform even worse than the driving

**Fig. 4** Histogram and quantiles (0.025, 0.1, 0.25, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95 and 0.99) of daily mean precipitation (mm/day) for the four seasons: DJF (December–January–February), MAM (March–April–May), JJA (June–July–August) and SON (September–October–November) as simulated with the WRF model (30 and 10 km resolutions), the original ERA-40 reanalysis, as observed and as simulated with the models included in the ENSEMBLES project



reanalysis in reproducing the quantiles from 0.6 to 0.99. The ENSEMBLES mean performs very well. The WRF model shows good skill and the refined resolution of 10 km adds more value to the 30 km simulation producing nearly a perfect agreement between the observed and modelled quantiles. We also see that these results are consistent during all seasons.

### 3.3.3 Regional differences of precipitation

Due to the long and narrow form and complex terrain of Norway the regional differences in precipitation can be large. In Hanssen-Bauer et al. 1997, 13 different regions have been defined to represent the geographical diversity of precipitation. These regions are shown in Fig. 5 with their

observed accumulated precipitation from 1961 to 1990 averaged over the stations in each region. The number of stations in each region varies from 4 to 50 depending on the size of the region. To calculate the regional means the stations in each region are pooled together for bias calculation, and then averaged. We use these regions to assess how well the WRF model is able to reproduce the precipitation in the different regions. Figure 5 illustrates the bias of total accumulated precipitation in 1961–1990 per region. We see that the regions 2, 4, 6–9 and 13, which are the wet regions along the coast, are generally well reproduced by the models. As was already seen in Fig. 2 the precipitation is overestimated by all models in the drier regions 3, 10 and 12. The precipitation in these regions is largely influenced by summertime convection which often causes large biases between model and observations. Both WRF simulations are performing quite well compared with other models. In average the 30-km simulation is overestimating the precipitation slightly less than the 10-km simulation and both of them have slightly better skill than the ENSEMBLES mean.

This is also the case for the extreme precipitation which is defined as 0.95 quantile. The figure did not change if we changed the 0.95 quantile to 0.9, 0.99 or 0.999. In relative numbers the models mostly overestimate the extreme precipitation in the driest areas (3 and 10–12) and perform

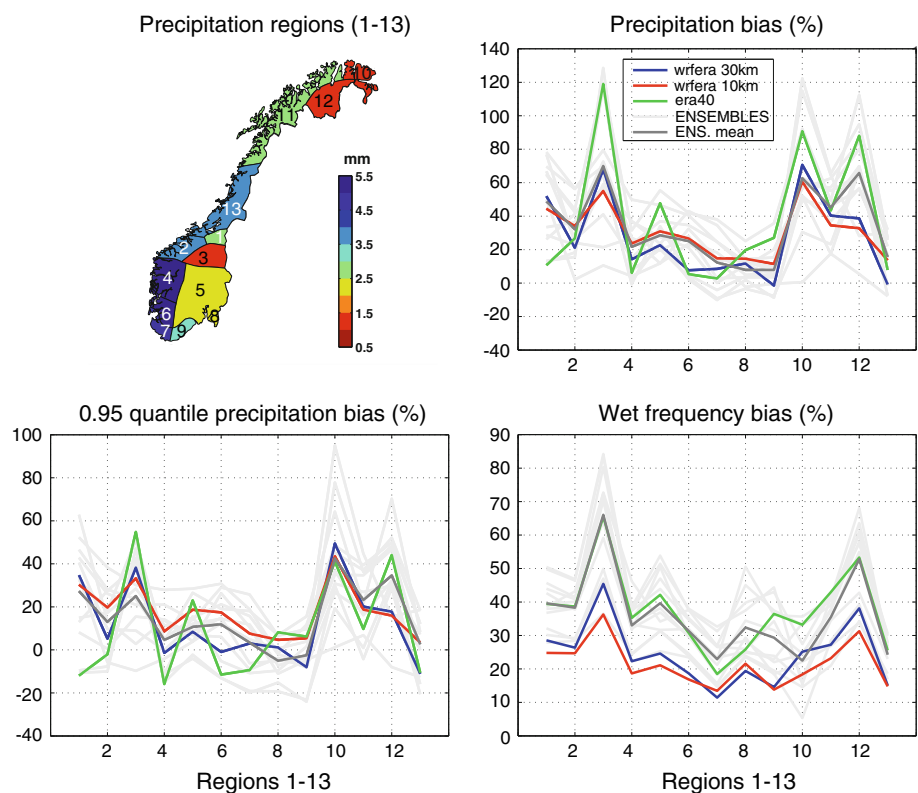
better along the wet west coast of Norway. Both WRF simulations are comparable with the ENSEMBLES mean.

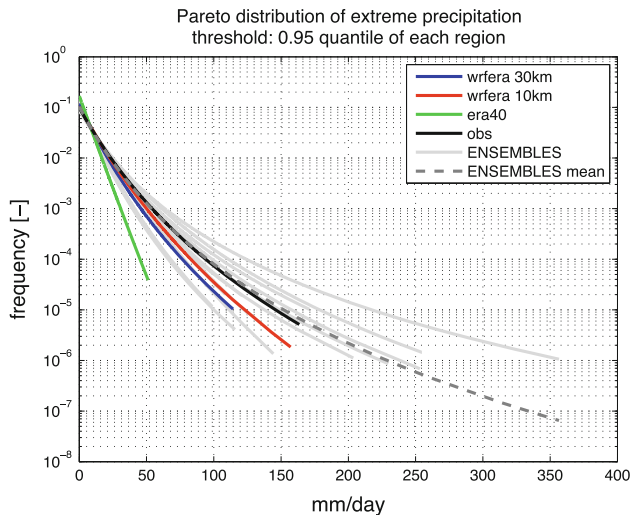
The bias in the number of wet days (Fig. 5) shows that all models overestimate the frequency of precipitation. The error is largest again in the dry areas, 3 and 12, but otherwise the difference between the regions is smaller than in the case of total precipitation of extreme precipitation. The WRF model seems to be performing very well giving a low bias compared with the ENSEMBLES models.

### 3.3.4 Extreme values of precipitation

The high precipitation events in Norway are often connected with hazardous hydrological consequences, floods, landslides and the like. Therefore it is of special importance for the WRF simulations to reproduce the correct form of the higher end of the precipitation spectrum. We investigate the skill of the model using the generalized extreme value distributions (Coles 2001) for excesses of precipitation following the work of Coelho et al. 2007. Excesses are defined as exceedances of precipitation over a certain threshold. Defining this threshold is not straightforward because it has to be high enough to describe an extreme value but still a large enough number of values must be higher than that so that statistical significance is reached. As discussed in the previous section the amount of

**Fig. 5** The 13 precipitation regions of Norway with their respective observed mean precipitation in 1961–1990 in mm/day, precipitation bias (%) per region, bias of the 0.95 quantile (%) per region and bias of frequency of wet days (%) per region





**Fig. 6** Generalized extreme value distribution of high precipitation, defined as 0.95 quantile of daily mean precipitation in each region separately

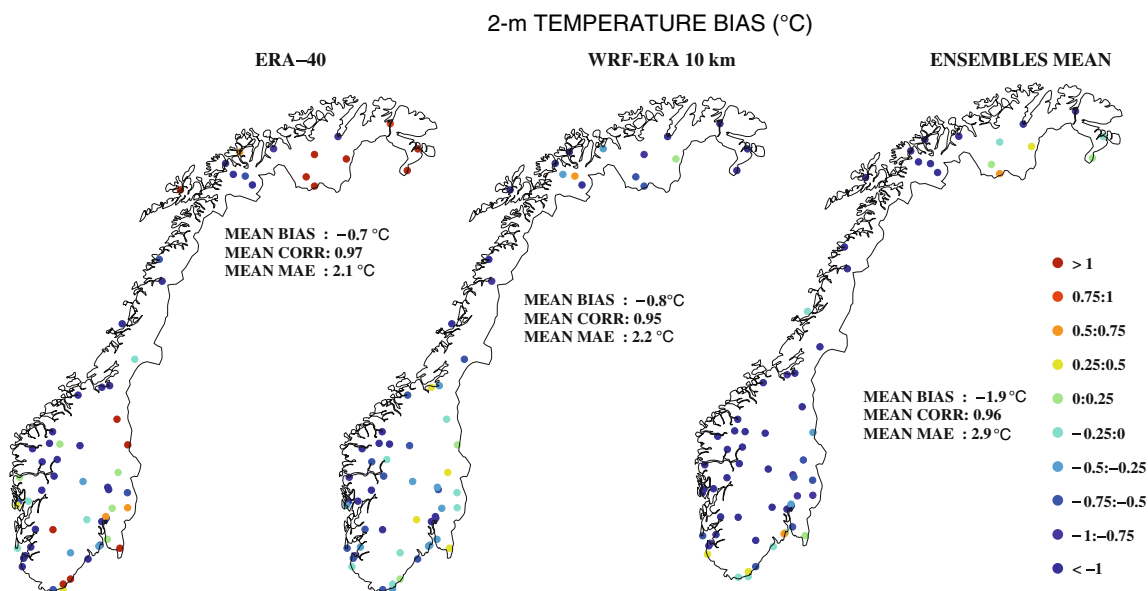
precipitation varies largely between the precipitation regions in Norway. A certain amount of precipitation might be dangerous in a dry region but represent only an average amount in wetter regions. To account for these differences we define the threshold to be the 0.95 quantile precipitation for each region, calculate the exceedances for each region separately and then fit that data to a generalized extreme value distribution (Fig. 6). The threshold ranged from 11 to 37 mm/day between the regions. We see from the figure that the ERA-40 reanalysis is not reproducing the extreme values of precipitation due to its coarse horizontal

resolution. Increasing the resolution to 10–30 km improves the representation of extreme values clearly (WRF 30 km, WRF 10 km and the ENSEMBLES models). Still, there is a large spread between the models. We see that increasing the resolution of the WRF model improves the results and that the WRF model is performing generally well compared with the individual ENSEMBLES models. Some of the ENSEMBLES models give enormous extreme values as was already seen in the histogram (Fig. 4) and these values are distorting the extreme value distribution. The form of the distribution of the ENSEMBLES mean is almost perfect, but the upper end of the spectrum is largely overestimated.

### 3.4 2-m temperature

#### 3.4.1 Geographical distribution of 2-m temperature bias

The geographical distribution of the 2-m temperature bias calculated from the daily mean values between the model and observations is illustrated in Fig. 7. The average statistics of the ERA-40 and WRF 10-km simulation show no significant difference. Both the reanalysis and the WRF model predict a cold bias of 0.7–0.8°C over the country. The WRF run is reducing the warm bias in the northern Norway as well as on the south coast. The mean of the ENSEMBLES models is performing slightly worse—the mean bias (−1.4°C) and the mean absolute error (2.7°C) are half-a-degree larger than that of the ERA-40 or WRF 10-km simulation. The mean correlation coefficients calculated from the daily mean values are very good in all



**Fig. 7** The 30-year mean 2-m temperature bias, ERA-40 reanalysis, simulated with the WRF model (10 km) and the 12 model mean of the ENSEMBLES project



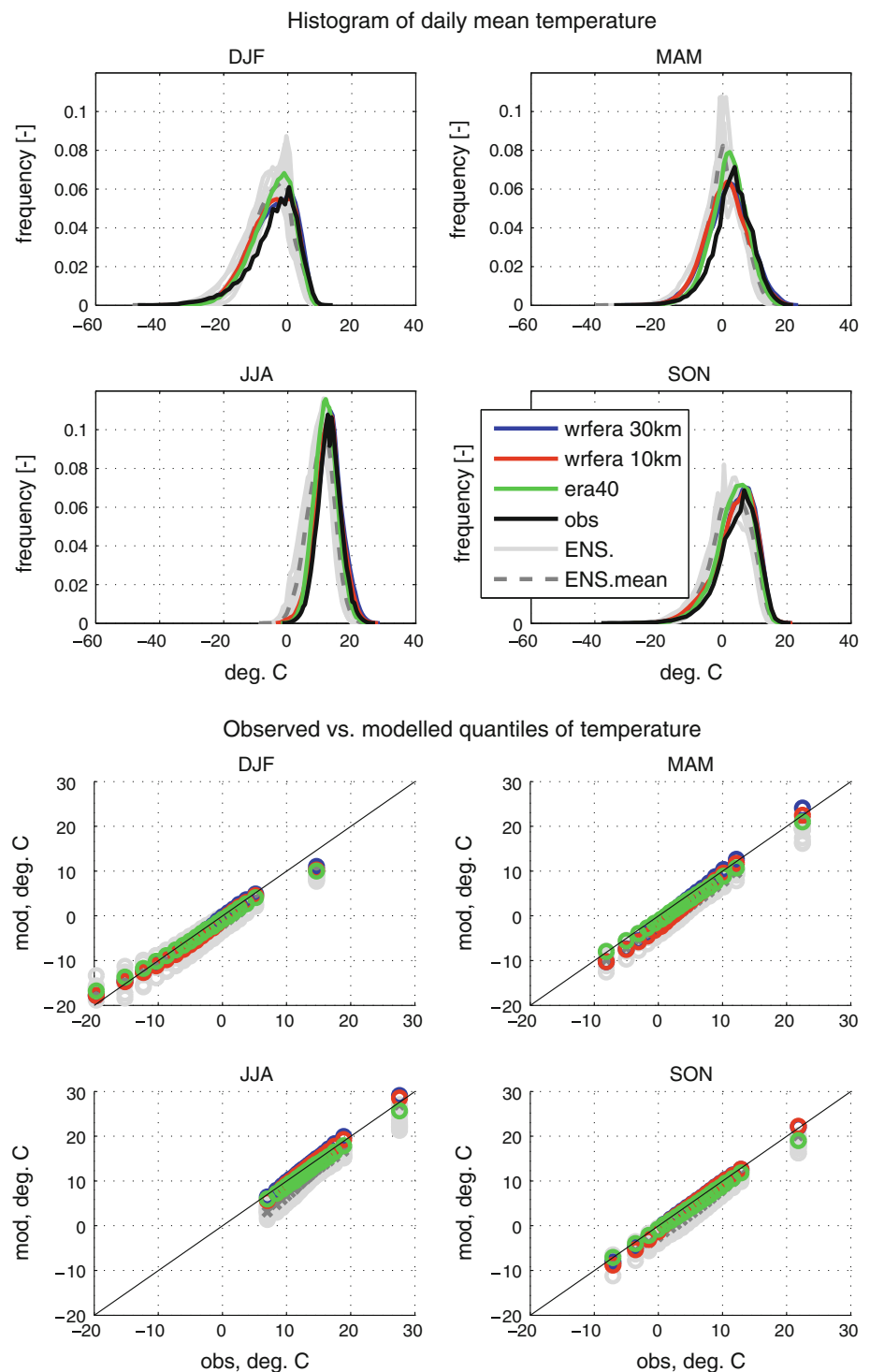
cases (0.95–0.97), partly due to the reproduction of the seasonal cycle. The ERA-40 temperature bias has a strong east-west gradient, the temperatures being too cold (ca.  $-1^{\circ}\text{C}$ ) on the west coast and too warm (ca.  $1^{\circ}\text{C}$ ) in the eastern part. This gradient is inherited by the ENSEMBLES models and their mean, but reduced in the WRF 10 km simulation. We argue that this could be caused by

the large outer domain used, giving the WRF model more freedom to deviate from the driving data.

### 3.4.2 Histogram of the daily mean temperature

The upper part of Fig. 8 shows the histogram of the daily mean temperature values of the observations, ERA-40

**Fig. 8** Histogram and quantiles (from 0.05 to 1 in steps of 0.05) of the daily mean 2-m temperature ( $^{\circ}\text{C}$ ) as simulated with the WRF model (30 and 10 km resolutions), the original ERA-40 reanalysis, as observed and as simulated with the models included in the ENSEMBLES project



reanalysis, 30- and 10-km WRF simulations, the individual ENSEMBLES models and the ENSEMBLES mean for the four seasons. The histograms of all ENSEMBLES models, and consequently the ENSEMBLES mean, are shifted towards cold temperatures during all seasons compared with the observations. The ERA-40 and the WRF simulations perform well during the summer but share the cold bias during other seasons. The coldest observed temperatures (between  $-30$  and  $-50^{\circ}\text{C}$ ) during the DJF and partly MAM seasons are missing in all model data. These cold extremes occur on clear sky winter days with strong inversions which are not well simulated by models. This is a general problem in numerical weather prediction (e.g. Mölders and Kramm 2010; Tjernström et al. 2004). This causes the cold temperatures of the histogram to be shifted towards milder temperatures ( $0$  to  $-20^{\circ}\text{C}$ ). The upper end of the histogram is close to the observed.

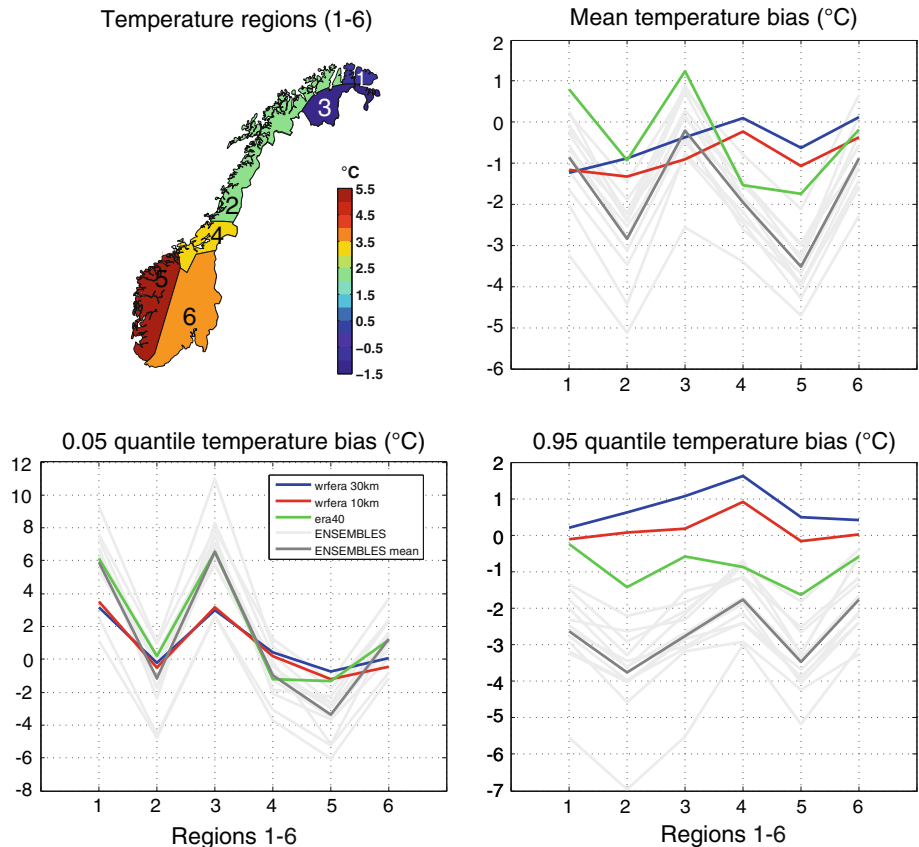
The same is shown by the modelled quantiles in the lower part of Fig. 8, plotted against the observed quantiles. Both WRF simulations reproduce well the higher quantiles but overestimate the lower quantiles during the DJF season. The results of the 10-km nest are slightly improving the 30-km results of the WRF model. There is a large spread between the individual ENSEMBLES models in the lower end of the quantiles but the ENSEMBLES mean is performing quite well. In the upper end of the quantiles all

ENSEMBLES models underestimate the observed temperatures leading to a larger general cold bias than the bias of the WRF simulations.

### 3.4.3 Regional differences of temperature

Similarly to the precipitation regions discussed in Sect. 3.3.3, the Norwegian meteorological office has defined 6 different temperature regions (Hanssen-Bauer and Førland 2000). These regions can be seen in Fig. 9 with their 30-year mean observed temperatures. The west coast is the warmest region (5), followed by the eastern part of the country (6). Average temperatures drop the further north the regions are located, with the region 3 in the northern inland as the coldest. The regional biases are calculated the same way as in the case of precipitation. The WRF simulations are outperforming the ENSEMBLES models when looking at the regional mean temperature bias (upper right panel of the Fig. 9). The WRF simulations reproduce the regional differences quite well with best agreement in the regions 4 and 6 and the weakest agreement in the regions 1 and 2. The 30-km WRF simulation performs better than the 10-km simulation in all regions with an average difference of almost  $0.5^{\circ}\text{C}$ . The ENSEMBLES mean is underestimating the temperature by up to  $2\text{--}3$  degrees in the regions 2, 4 and 5 but performs well in other regions. The ERA-40

**Fig. 9** The six temperature regions of Norway with their respective observed mean temperature in 1961–1990 in  $^{\circ}\text{C}$ , mean temperature bias ( $^{\circ}\text{C}$ ) per region, bias of the 0.05 quantile temperature ( $^{\circ}\text{C}$ ) per region and bias of the 0.95 quantile temperature ( $^{\circ}\text{C}$ ) per region



reanalysis has large positive and negative biases in different regions which compensate each other giving a reasonably small mean bias. Comparing the WRF simulations and the ENSEMBLES models with the driving ERA-40 data we see that the ENSEMBLES models all have same cold and warm biases as the ERA-40 but are generally colder. The WRF model, instead, seems to be more independent of the driving data and is able to keep the bias low in all regions. This could be due to the larger domain size used in the WRF simulation indicating that it is an asset and also the spectral nudging procedure used above the boundary layer.

The lower two panels of Fig. 9 show the 0.05 and 0.95 quantile temperatures describing the extremely low and extremely high temperatures. Both WRF simulations are in good agreement with the observed extreme temperatures and outperform the ENSEMBLES mean or the ERA-40 reanalysis. There are no large differences between the regions in the extremely high temperatures and the 10-km WRF simulation is giving the best results. The models vary more in reproducing the extreme low temperatures. All models and the ENSEMBLES mean perform well in the southern regions 4–6 but fail to reproduce the extremely cold temperatures of below  $-30^{\circ}\text{C}$  observed in the northern regions, as discussed in the previous section.

#### 3.4.4 Extreme values of temperature

The analysis was performed in the same way as described in Sect. 3.3.4 for the upper and the lower end of the temperature spectrum. The generalized extreme value distributions are shown in Fig. 10 for the extreme low (threshold: 0.05 quantile of each region; between  $-23$  and  $-6^{\circ}\text{C}$ ) and extremely high (threshold: 0.95 quantile of each region; between  $14$  and  $17^{\circ}\text{C}$ ) temperature. The  $x$ -axis in the figures shows the exceedances from the threshold in

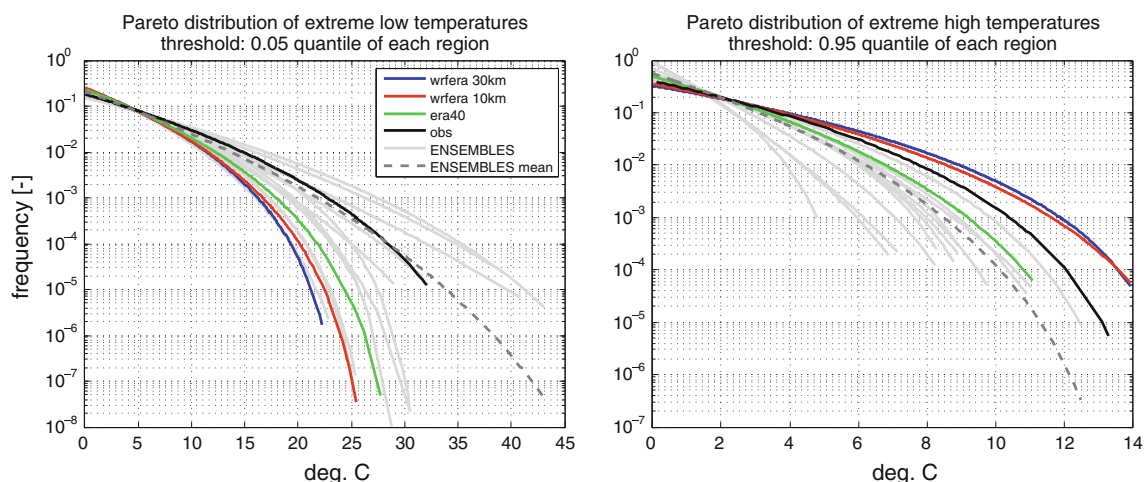
magnitude, not an absolute temperature. We see that the spread of the excesses of the extremely low temperatures ( $0$ – $30^{\circ}\text{C}$ ) is much larger than that of extremely high temperatures ( $0$ – $12^{\circ}\text{C}$ ) which presents an additional challenge for the models.

Generally the agreement between the modelled and observed excesses is satisfying. The WRF simulations fail to reproduce the extremely low temperatures as we have seen in Fig. 8. The 10-km simulation is improved from the 30-km simulation but still lacking the extremely cold temperatures. The ERA-40 and all of the ENSEMBLES models have a better agreement with the observed distribution than the WRF simulations. A few of the ENSEMBLES models produce too cold extreme temperatures and give the distribution of the ENSEMBLES mean too long a tail but the shape of the distribution is correct.

The situation is changed in the case of the extremely high temperatures. The WRF simulations are reproducing the distribution of the observed temperatures reasonably well. There is almost no difference between the 30 and 10-km results of the WRF model. The WRF model is overestimating the 0.95 quantile temperatures whereas the ENSEMBLES models are underestimating them. The error on both sides is approximately as large. These differences reflect the overall shift towards cold temperatures of the ENSEMBLES models compared with the WRF simulations.

#### 3.5 10-m wind speed

The winds are generally well simulated or slightly too low (in the order of  $1$ – $2$  m/s) on the coast and overestimated (up to  $>50\%$ ) in the inland stations in all models (not shown). The mean statistics show that all models are very similar and that refining the horizontal resolution from 30 to 10 km

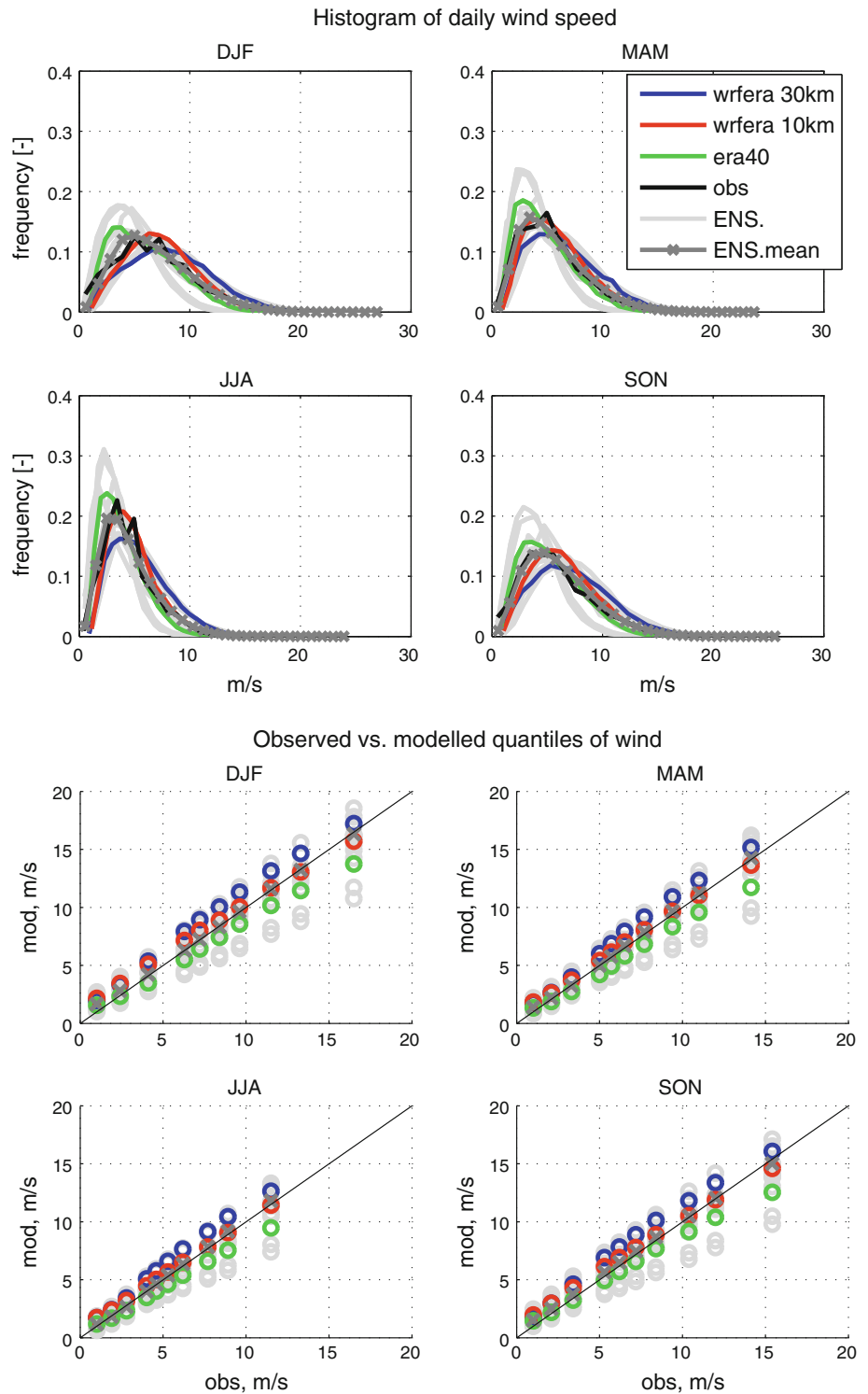


**Fig. 10** Generalized extreme value distribution of low (0.05 quantile of each region) and high (0.95 quantile of each region) daily mean temperatures

does not make a significant difference. This is likely to be due to the land use data used in the models, which generally does not describe the Norwegian vegetation in high detail. In studies which concentrate on surface winds a higher horizontal resolution as well as use of a more detailed description of land use would be important.

The histogram in the upper panel of Fig. 11 illustrates the daily mean wind speed of the ten coastal stations mentioned before for the four seasons. A comparison shows that almost all models overestimate the low winds but underestimate the high winds, except the ERA-40 reanalysis which mainly underestimates the winds during

**Fig. 11** Histogram of daily mean 10-meter wind speed (m/s) and quantiles as simulated with the WRF model (30 and 10 km resolutions), the ERA-40 reanalysis, as observed and as simulated with the models included in the ENSEMBLES project. Only ten representative coastal stations are included



all seasons. The 10-km WRF simulation is clearly improving the 30-km simulation reducing the windy bias similarly to the ENSEMBLES mean, especially in the low end of the spectrum. This is likely to be due to the more realistic representation of the coastline in the 10-km grid.

The quantiles in the lower panel of Fig. 11 show that despite of the over- (under-) estimation of the low (high) winds the form of the histogram is reasonable. Only the lower quantiles (from 0 to 0.5) are significantly overestimated in the 10-km WRF simulation. Here we see a clear improvement of the 10-km nest in the WRF simulation. The spread of the ENSEMBLES models is large but the ENSEMBLES mean agrees very well with the observed quantiles.

#### 4 Summary and discussion

Results are presented from a dynamical downscaling of the ERA-40 reanalysis, with the WRFV3.1.1 regional climate model, to 30 and 10 km resolutions for 1961–1990 in Norway. The results of 12 different regional climate model simulations from the ENSEMBLES project are also presented as a reference. We concentrate the analysis on surface variables on complex terrain: precipitation, 2-m temperature and 10-m wind speed and compare the model results with a large number of observations within Norway.

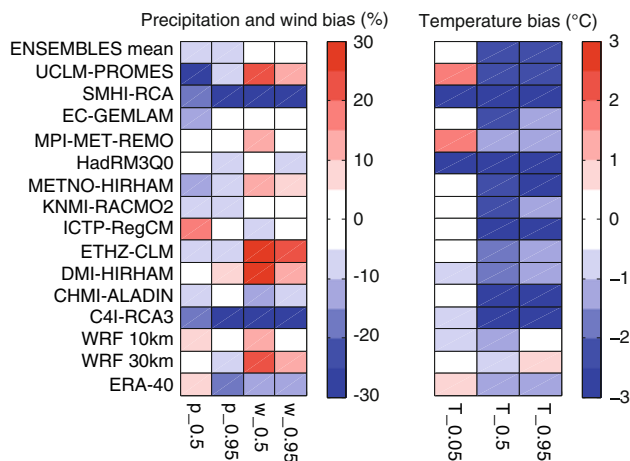
Figure 12 summarizes the general behaviour of all experiments analyzed within this study. The biases shown are deviations of the daily mean modeled values of the observations, averaged over all stations over the whole period of 1961–1990. Precipitation and wind biases are

shown in percent and the temperature biases in degrees. We focus on the “mean” (0.5 quantile) and “extreme” (0.95 quantile for extremely high and 0.05 quantile for extremely low) values. The figure shows that there is large spread in the quality of the modeled precipitation and wind between the individual models. The WRF simulations perform comparably well and the value added by the refinement of the resolution to 10 km is obvious. The ENSEMBLES mean has low biases and only a few of the models are performing better. In case of temperature the WRF simulations have clearly lower biases than the individual ENSEMBLES models or the ENSEMBLES mean. Again, the 10-km simulation reduces the bias compared with the 30-km simulation.

The precipitation on the Norwegian coast is largely driven by advective systems. As opposed to the traditional setup for regional climate models downscaling the European climate we included the whole North Atlantic into the larger model domain and applied spectral nudging to keep the large scale circulation consistent with the driving data. This turned out to be advantageous in several ways. First, the phase of the precipitation events was improved from the ERA-40 indicating that the synoptic scale features were better resolved by the 30 km grid than in the reanalysis. Also the representation of extreme precipitation on the coast was much improved from the reanalysis, probably due to sharper gradients and better resolved fronts. Another advantage seemed to be the larger independence of the regional model compared with the driving data. The WRF simulations were able to reduce the large regional biases of surface temperature in the ERA-40 reanalysis which had been largely inherited by the ENSEMBLES simulations.

A relatively high horizontal resolution turned out to be important in complex terrain, such as the Norwegian coast and the mountains. The precipitation has a large orographic enhancement which was largely improved from the reanalysis by the WRF simulations. The orographic lifting in the 10-km simulation was stronger and better resolved than in the 30-km simulation which also lead to an improvement of the representation of the extreme precipitation events, especially in the mountains. We conclude that the use of a horizontal resolution of 10 km, or higher, is preferable for producing climate projections, especially for impact studies dealing with extreme precipitation.

The fact that the precipitation and coastal winds are improved on a higher resolution grid is a consequence of a better representation of topography and coastline. This is in accordance with the general findings from several regional climate model studies (Rummukainen 2010). It also has to be kept in mind that precipitation of the ERA-40 reanalysis is a pure model product but temperature and winds are more constrained by the observations which improves the agreement. Also the fact that temperature and winds from



**Fig. 12** Models used for analysis in this study which participated in the ENSEMBLES project. The horizontal resolution of all models was 25 km and the years included in the analysis were 1961–1990. The figure shows the 30-year mean model bias of median ( $X_{0.5}$ ) and 0.95-quantile ( $X_{0.95}$ ) precipitation and wind (%) and the bias of 0.05-quantile, median and 0.95-quantile temperature ( $^{\circ}\text{C}$ ) averaged of the daily mean values of each station separately



the ERA-40 reanalysis are input fields for the WRF model, but precipitation not, explains why the differences between the temperature and winds of WRF runs and ERA-40 are smaller than for precipitation.

This study was the first application of the WRF model to climatological simulations in Europe. Generally the WRF model performed very well in reproducing the observed climate in Norway. The default setup of physical schemes in the WRF model turned out to be a suitable approach in climatological studies keeping the runtime low but producing results similar to the more sophisticated schemes. Spectral nudging proved to be a very useful method in these simulations where the outer model domain was large. The phase of precipitation and temperature was significantly improved in the nudged runs compared with the free runs (not discussed in this paper) and the simulated extreme values of precipitation were more realistic. This model configuration is useful for downscaling of GCM future predictions and the high-resolution data set created provides input for further downscaling and impact studies.

**Acknowledgments** This work has been financed by the Norwegian Research Council through the SFF grant to the Bjerknes Centre for Climate Research and by the Municipality of Bergen through the MARE project. The authors acknowledge the computer resources and support by the Parallab at the Uni Research in Bergen, Norway. The ENSEMBLES data used in this work was funded by the EU FP6 integrated project ENSEMBLES (contract number 505539) whose support is gratefully acknowledged. This is publication no. A-307 from the Bjerknes Centre for Climate Research.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

- Alexandru A, de Elia R, Laprise R, Separovic L, Biner S (2008) Sensitivity study of regional climate model simulations to large-scale nudging parameters. *Mon Wea Rev* 137:1666–1686. doi:[10.1175/2008MWR2620.1](https://doi.org/10.1175/2008MWR2620.1)
- Barstad I, Sorteberg A, Flatøy F, Déqué M (2009) Precipitation, temperature and wind in Norway: dynamical downscaling of ERA40. *Clim Dyn* 33:769–776. doi:[10.1007/s00382-008-0476-5](https://doi.org/10.1007/s00382-008-0476-5)
- Coelho CAS, Ferro CAT, Stephenson DB, Steinskog DJ (2007) Methods for exploring spatial and temporal variability of extreme events in climate data. *J Clim* 21:2072–2092
- Coles S (2001) An introduction to statistical modeling of extreme values. Springer series in statistics, Springer, Berlin
- Collins WD, Rasch PJ, Boville BA, Hack JJ, McCaa JR, Williamson DL, Briegleb BP, Bitz CM, Lin SJ, Zhang M (2006) The formulation and atmospheric simulation of the community atmosphere model version 3 (CAM3). *J Clim* 19(11):2144–2161. doi:[10.1175/JCLI3760.1](https://doi.org/10.1175/JCLI3760.1)
- Ek MB, Mitchell KE, Lin Y, Rogers E, Grunmann P, Koren V, Gayno G, Tarpley JD (2003) Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. *J Geophys Res* 108(D22):8851. doi:[10.1029/2002JD003296](https://doi.org/10.1029/2002JD003296)
- ENSEMBLES members (2009) Climate change and its impacts at seasonal, decadal and centennial timescales. Summary of research and results from the ENSEMBLES project. Available at [http://ensembles-eu.metoffice.com/docs/Ensembles\\_final\\_report\\_Nov09.pdf](http://ensembles-eu.metoffice.com/docs/Ensembles_final_report_Nov09.pdf)
- Hanssen-Bauer I, Førland EJ, Tveito OE, Nordli PØ (1997) Estimating regional precipitation trends: comparison of two methods. *Nordic Hydrol* 28:21–36
- Hanssen-Bauer I, Førland E (2000) Temperature and precipitation variations in Norway 1900–1994 and their links to atmospheric circulation. *Int J Climatol* 20:1693–1708
- Hong S-Y, Dudhia J, Chen S-H (2004) A revised approach to ice microphysical processes for the bulk parameterization of clouds and precipitation. *Mon Wea Rev* 132:103–120
- Hong SY, Noh Y, Dudhia J (2006) A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon Wea Rev* 134:2318–2341
- Kain JS (2004) The Kain–Fritsch convective parameterization: an update. *J Appl Meteor* 43:170–181
- Jones RG, Murphy JM, Noguer M (1995) Simulation of climate change over Europe using a nested regional climate model. I: Assessment of control climate, including sensitivity to location of lateral boundaries. *Q J R Meteorol Soc* 121:1413–1449
- Køltzow M, Iversen T, Haugen JE (2008) Extended Big-Brother experiments: the role of lateral boundary data quality and size of integration domain in regional climate modelling. *Tellus A* 60(3):398–410. doi:[10.1111/j.1600-0870.2008.00309](https://doi.org/10.1111/j.1600-0870.2008.00309)
- Kostopoulou E, Tolika K, Tegoulas I, Giannakopoulos C, Somot S, Anagnostopoulou C, Maheras P (2009) Evaluation of a regional climate model using in situ temperature observations over the Balkan Peninsula. *Tellus A* 61(3):357–370
- Leung LR, Ghan SJ (1995) A subgrid parameterization of orographic precipitation. *Theor Appl Climatol* 52:95–118
- Míguez-Macho G, Stenchikov GL, Robock A (2004) Spectral nudging to eliminate the effects of domain position and geometry in regional climate model simulations. *J Geophys Res* 109:D13104. doi:[10.1029/2003JD004495](https://doi.org/10.1029/2003JD004495)
- Míguez-Macho G, Stenchikov GL, Robock A (2005) Regional climate simulations over North America: Interaction of local processes with improved large-scale flow. *J Clim* 18:1227–1246
- Mölders N, Kramm G (2010) A case study on wintertime inversions in Interior Alaska with WRF. *Atmos Res* 95(2–3):314–332
- Radu R, Déqué M, Somot S (2008) Spectral nudging in a spectral regional climate model. *Tellus* 60A:898–910
- Rummukainen M (2010) State-of-the-art with regional climate models. *Wiley Interdisciplinary Reviews: Clim Change* 1:82–96
- Stohl A, Forster C, Sodemann H (2008) Remote sources of water vapor forming precipitation on the Norwegian west coast at 60°N—a tale of hurricanes and atmospheric river. *J Geophys Res* 113:D05102. doi:[10.1029/2007JD009006](https://doi.org/10.1029/2007JD009006)
- Von Storch H, Langenberg H, Feser F (2000) A spectral nudging technique for dynamical downscaling purposes. *Mon Wea Rev* 128:3664–3673
- Tjernström M, Zagar M, Svensson G (2004) Model simulations of the Arctic atmospheric boundary layer from the SHEBA year. *AMBIO: A J Hum Environ* 33(4):221–227
- Uppala SM et al. (2005) The ERA-40 re-analysis. *Q J R Meteorol Soc* 131(612):2961–3012
- Waldron KM, Peagle J, Horel JD (1996) Sensitivity of a spectrally filtered and nudged limited area model to outer model options. *Mon Wea Rev* 124:529–547
- Zahn M, Von Storch H, Bakan S (2008) Climate mode simulation of North Atlantic polar lows in a limited area model. *Tellus* 60A:620–631